

La demenza senile: valutazione da un punto di vista statistico

Vittorio Bomba, Federico Luzzi, Marco Peracchi, Christian Uccheddu

Sommario

La demenza senile è un disturbo delle funzioni intellettive legate alla memoria, al linguaggio e al pensiero che colpisce individui in età avanzata, ovvero soggetti con un'età superiore ai sessant'anni. Tale malattia accoglie al suo interno diverse sfumature, che tengono conto della sua natura e delle sue cause, di seguito ne verranno indicate alcune:

- Demenza degenerativa, che include l'Alzheimer, attribuibile a difetti genetici
- Demenza vascolare, derivante dal susseguirsi di microinfarti ripetuti che distruggono il tessuto cerebrale
- Demenza reversibile, derivante da disfunzioni che possono essere corrette

Data la definizione molto generica e qualitativa di demenza[1] non esistono parametri clinici che possano univocamente certificare il disturbo, che viene attestato dal medico sulla base di esami medici e parametri ambientali. La necessità di valutare accuratamente la serietà di questa disfunzione ha portato alla creazione di un test, noto come CDR[2] (Clinical Dementia Rating), una valutazione numerica assegnata dal medico che riporta la gravità della demenza sulla base di un'intervista strutturata per verificare il deficit della memoria del paziente.

L'obiettivo del progetto è stimare il valore del CDR sulla base di modelli di machine learning, questo potrebbe essere utile per sostituire l'intervista al paziente o per fungere da supporto al medico nell'assegnazione del valore CDR.

Keywords: *machine learning, demenza, CDR*

Indice

1	Premessa	2
2	Descrizione Dataset	2
3	Data pre-processing	2
3.1	Variables selection	2
3.2	Missing values	3
3.3	Correlazione fra le variabili	3
4	Modelli	3
4.1	Tecniche Euristiche	3
4.2	Tecniche regression based	3
4.3	Tecniche di separazione	4
4.4	Tecniche probabilistiche	4
5	Elaborazione e interpretazione dei dati	4
5.1	Holdout	5
5.2	Iterated Holdout	5
5.3	Cross-validation	5
5.4	Feature selection	6
6	Conclusioni	7

1 Premessa

Il dataset oggetto del nostro studio si chiama *MRI and Alzheimers*[3], preso dalla piattaforma online di condivisione dati Kaggle[4], in particolare il file *oasis_longitudinal.csv*. Il campionamento effettuato è di tipo longitudinale: uno studio condotto in un determinato tempo, prendendo una porzione di popolazione. Esso raccoglie i dati clinici di 150 soggetti di età compresa tra i 60 e i 96 anni con una demenza senile *sospetta*. Ciascun record rappresenta una visita effettuata su un soggetto, almeno 2 visite per soggetto.

2 Descrizione Dataset

Il dataset è composto da 373 record organizzati in 15 colonne, ognuna delle quali esplicita un attributo della tupla. Di seguito la loro descrizione:

- **Subject ID**: sequenza di caratteri alfanumerici che identifica il paziente
- **MRI ID**: numero identificativo della risonanza magnetica eseguita sul paziente. Compare come il Subject ID seguito dal numero della risonanza magnetica
- **Group**: gruppo di appartenenza del paziente. Può essere "Demented", "Non Demented" o "Converted"
- **Visit**: numero della visita dell'individuo
- **MR delay**: tempo intercorso in secondi tra l'iniezione del liquido di contrasto e l'esecuzione della risonanza
- **M/F**: genere del paziente (maschio o femmina)
- **Hand**: mano dominante del soggetto (tutti i soggetti interessati usano la mano destra)
- **Age**: età del paziente
- **EDUC**: anni di istruzione
- **SES**[5]: status socioeconomico assegnato dall'Hollingshead Index of Social Position sulla base di un punteggio che va da 1 (più alto status) a 5 (più basso status)
- **MMSE**[6]: Mini-Mental State Examination, un test neuropsicologico costituito da 30 domande utilizzate per verificare lo stato cognitivo. Va da 1 (peggior stato) a 30 (miglior stato). Con un punteggio inferiore a 24 il paziente presenta indebolimento cognitivo ma non necessariamente demenza;
- **eTIV**[7]: estimated Intracranial Volume, stima del volume intracranico in cm^3

- **nWBV**: normalized Whole-Brain Volume, espresso come percentuale del volume del cervello in rapporto con il volume intracranico (eTIV)
- **ASF**: Atlas Scaling Factor[8], determinante della matrice di trasformazione tra lo spazio della MRI e la mappa utilizzata per strutturare il cervello. In pratica restituisce il rapporto fra la misura del volume dell'immagine e della mappa.
- **CDR**[2]: Clinical Dementia Rating, scala numerica[9] che identifica la gravità della demenza. In particolare si ha 0 per un paziente normale, 0.5 per dubbia demenza, 1 per lieve demenza, 2 per moderata demenza, 3 per grave demenza. Infine 4 e 5 rappresentano i casi più gravi in assoluto.

Il dataset usato presenta la seguente distribuzione di CDR:

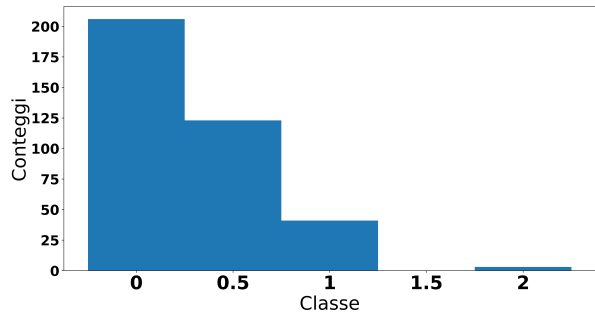


Figura 1: Le classi più basse sono molto più presenti nel dataset, la classe 2 si presenta con soli tre record mentre le restanti non sono presenti

3 Data pre-processing

Nella fase di data-preprocessing il lavoro svolto è stato quello di:

1. Escludere le variabili inutili per il modello
2. Riempire i valori delle celle mancanti
3. Eliminare le correlazioni fra le variabili

3.1 Variables selection

Gli attributi esclusi sono stati:

- **Subject ID**, **MRI ID** in quanto stringhe alfanumeriche identificative del paziente. La variabile **Hand** è stata esclusa in quanto i soggetti sono tutti destrorsi.
- **Visit**, indica il numero della visita in cui si è stabilito il grado di demenza del paziente, rappresentando come "pazienti diversi" visite differenti. Questa esclusione potrebbe portare ad errori nel modello predittivo, tuttavia nella fase della data exploration ci

siamo accorti che alcuni pazienti precedentemente giudicati dementi risultavano nelle visite successive non dementi. Di conseguenza si è cercato di evitare che il numero della visita influenzasse il modello in quanto le uniche sostanziali variazioni tra le visite riguardano l'età del paziente.

- **Group**, che presenta 3 valori: *Demented*, *NonDemented* e *Converted*, quest'ultimo è di difficile interpretazione poiché il paziente era prima considerato *NonDemented* e poi *Demented* o viceversa. Inoltre abbiamo deciso di escludere questo attributo in quanto rappresenta la stessa informazione della nostra variabile di classe CDR e infine si è preferito decidere di considerare CDR piuttosto che Group siccome più completa, esplicativa ed utile al medico nella definizione di persona con demenza.

3.2 Missing values

Il dataset presenta 19 valori mancanti alla voce SES e 2 alla voce MMSE. La trattazione di tali valori è stata affrontata in modo differente:

- **SES**: i valori mancanti sono stati sostituiti dall'interpolazione lineare con i dati presenti dell'attributo in questione. Questo per mantenere la distribuzione dei dati piuttosto omogenea e al contempo non dover eliminare 19 record su 373.
- **MMSE**: le righe corrispondenti ai valori mancanti sono state scartate, in quanto tale valore non è ottenibile osservando gli altri attributi o la distribuzione dello stesso.

3.3 Correlazione fra le variabili

Alcuni attributi del dataset risultano essere legati tra di loro. Vista la definizione di eTIV, nWBV e ASF un legame tra di essi sembra suggerito. Analizzando gli attributi si può notare una correlazione quasi lineare tra ASF e il volume intracranico, in quanto all'aumentare del volume intracranico diminuisce il fattore di scala. La spiegazione risiede nel fatto che aumentando il volume intracranico aumenta, di conseguenza, anche il possibile volume del cervello, causando quindi la diminuzione del fattore di scala tra il cervello del paziente e quello standard, ovvero l'ASF. La percentuale di volume cerebrale (nWBV), invece, non risulta avere connessioni con l'ASF, questa percentuale è infatti insensibile alle variazioni di grandezza del cranio.

Le precedenti considerazioni ci portano a **escludere** l'attributo ASF dai modelli.

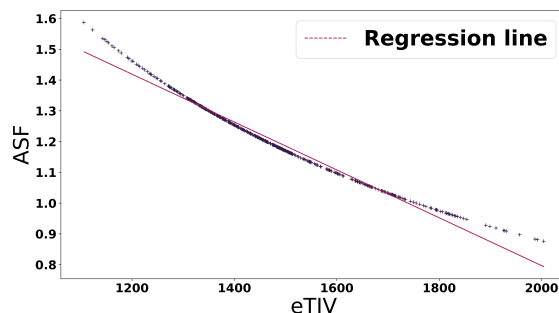


Figura 2: Il grafico mostra sulle ascisse il volume intracranico (eTIV) e sulle ordinate l'Atlas Scale Factor (ASF), il coefficiente di correlazione lineare è: -0.989

4 Modelli

Data la natura discreta della variabile di target CDR si è deciso di affrontare la criticità come un problema di classificazione. Sono stati valutati diversi modelli appartenenti a varie tecniche di classificazione per identificare la più adatta a questo caso specifico. Di seguito vengono elencati i modelli provati.

- **Tecniche Euristiche** (decision tree): Random Forest, J48
- **Tecniche Regression Based**: Regressione Logistica
- **Tecniche di Separazione**: Multi Layer Perceptron
- **Tecniche Probabilistiche**: Naive Bayes, Naive Bayes Tree

4.1 Tecniche Euristiche

J48: è una tecnica basata sul concetto di albero decisionale, in cui la particolarità è quella di riuscire a classificare anche dati nominali. Nella nostra implementazione abbiamo impostato un fattore di confidenza di 0.25 e un numero di fold pari a 3.

Random Forest: La Random Forest è una tecnica basata su comitati di alberi decisionali. Nella nostra implementazione sono stati inseriti 10 alberi con un fattore di confidenza di 0.25 e 3 folds.

4.2 Tecniche regression based

Regressione logistica: è una tecnica di classificazione basata sulla regressione, viene infatti calcolata la probabilità a posteriori che la variabile target restituisca il valore di input. La formula usata dal nodo Knime per calcolare la probabilità associata alla classe j-esima, a meno dell'ultima

classe, è la seguente:

$$P_j(X_i) = \frac{e^{X_i B_j}}{\sum_{j=1}^{k-1} e^{X_i B_j} + 1}$$

Mentre la probabilità associata all'ultima classe è espressa come:

$$1 - \sum_{j=1}^{k-1} P_j(X_i)$$

Per trovare la regressione migliore è necessario minimizzare la funzione Likelihood relativa a questa distribuzione di probabilità.

4.3 Tecniche di separazione

MLP (Multi Layer Perceptron): è un modello di classificazione basato sulla separazione dello spazio degli attributi, più precisamente consiste in neuroni che comunicano in modo feedforward dalle variabili di input alla variabile di classe. In particolare è stata predisposta una rete a **2 strati nascosti con 5 neuroni per strato**. Questa scelta di parametri è dovuta ad una serie di prove empiriche effettuate per l'ottimizzazione della classificazione.

4.4 Tecniche probabilistiche

Naive Bayes: è una tecnica di classificazione basata sull'applicazione del teorema di Bayes; un classificatore di Bayes combina il modello con una regola decisionale. Questo usa la regola per assegnare alla variabile \hat{y} l'indice di classe C_k da predire. Nel metanodo inserito in Knime è stata usata la regola che massimizza la funzione che calcola la probabilità di ottenere la predizione $\hat{y} = C_k$ dato il record $\{x_i\}$, ovvero la **MAP decision rule**:

$$\hat{y} = \arg \max_{k \in \{k_1, k_2, \dots, k_n\}} (p(C_k) \prod_{i=1}^n p(x_i | C_k))$$

Naive Bayes tree: è un modello che genera un albero decisionale in cui è implementato un classificatore di Bayes in corrispondenza dei nodi.

5 Elaborazione e interpretazione dei dati

Per valutare i risultati sono stati utilizzati diversi indici, primo fra tutti l'**accuracy**. L'accuracy misura l'affidabilità del modello a predire risultati corretti su nuovi dati. Per calcolare l'accuracy sul dataset D , con n entrate, a predire la variabile y_i dati i valori $\{x_i\}$ ($f(x_i)$ rappresenta la previsione) si sfrutta la seguente formula:

$$acc(D) = 1 - \frac{1}{n} \sum_{j=1}^n L(y_i, f(x_i))$$

L rappresenta la *loss function*, funzione che assegna valore 1 se $y_i \neq f(x_i)$ (cioè la predizione è errata) e 0 viceversa. Si può notare quindi che all'aumentare delle previsione errate diminuisce il valore dell'accuracy. I modelli utilizzati restituiscono la matrice di confusione. Questa contiene al suo interno diversi valori, nel nostro caso i valori di CDR che potevano essere predetti erano 0, 0.5, 1 e 2. Un esempio di matrice di confusione appare così:

		CDR			
		0	0.5	1	2
CDR _{pred}	0	60	8	0	0
	0.5	12	18	6	0
	1	1	1	10	0
	2	0	1	0	0

Sulla diagonale si possono notare i valori previsti correttamente dal modello, mentre gli altri rappresentano valori previsti in maniera scorretta.

La accuracy è calcolata sommando i valori sulla diagonale e dividendoli per il totale.

Il secondo indice utilizzato è la **precision**, definita come il rapporto tra i valori predetti correttamente della variabile che consideriamo diviso il numero di valori totali che appaiono nel dataset per quella variabile. Per esempio la precision della variabile 0.5 è:

$$P(0.5) = \frac{18}{12 + 18 + 6} = 0.50$$

Un altro indice sfruttato è la **recall**, calcolata sfruttando di nuovo il numero di valori predetti correttamente per la variabile considerata diviso il numero di valori totali predetti dal modello per quella variabile. Nuovamente riproponiamo come esempio la variabile 0.5, per cui la recall sarà:

$$R(0.5) = \frac{18}{8 + 18 + 1 + 1} = 0.64$$

L'ultimo indice che abbiamo sfruttato è la **F-measure**[10], calcolata come la media armonica tra gli indici elencati precedentemente:

$$F = \frac{2 \cdot r \cdot p}{r + p}$$

Valori elevati della F-measure indicano valori elevati di entrambi recall e precision. Per concludere l'esempio portato precedentemente otteniamo:

$$F(0.5) = \frac{2 \cdot 0.50 \cdot 0.64}{0.50 + 0.64} = 0.56$$

Per calcolare il valore di un indice complessivo di ogni modello è stata calcolata la **media aritmetica** degli indici calcolati sulle singole classi di quel modello nel modo seguente:

$$I_{model} = \frac{I(0) + I(0.5) + I(1) + I(2)}{4}$$

Con I: accuracy, F-Measure, recall, precision.

In questo modo è possibile confrontare i modelli.

5.1 Holdout

La tecnica di Holdout consiste nel dividere il dataset in due partizioni usando un campionamento stratificato, ovvero mantenendo le proporzioni dell'attributo di classe delle partizioni uguali alle proporzioni dell'attributo nel dataset di partenza. Questo metodo permette di sfruttare una parte del dataset come *train* (67%), per addestrare il codice, e una parte come *test* (33%), per valutare le predizioni fatte. Le misure di performance vengono eseguite sulla partizione di test. La problematica principale di questa tecnica è la dipendenza dell'accuracy dal partizionamento iniziale. Determinate divisioni potrebbero sovrastimare o sottostimare l'accuracy.

Di seguito le accuracy rilevate nella partizione di training rispetto a quelle rilevate nella partizione test:

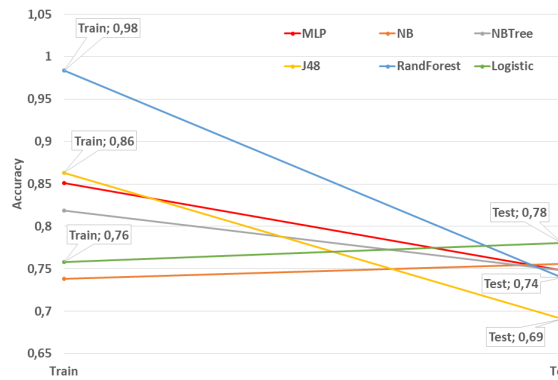


Figura 3: Il grafico mostra la variazione tra le accuracy di train (sinistra) e test (destra) per ogni modello

Come si può osservare dalla figura il modello Random Forest ottiene l'accuracy più alta sul dataset di training (0.98), ciò indica un *overfitting* del modello: infatti nel dataset di test si ottiene un punteggio nettamente inferiore (0.74). Il modello di regressione logistica si dimostra essere il più performante nel dataset di test, migliorando il suo punteggio rispetto al train. In generale tutti i modelli ottengono un buon livello di accuracy, attorno al 70%.

Nella tabella sottostante vengono mostrati i valori degli indici calcolati sulla partizione di test.

Modello	Recall	Precision	F-Measure	Accuracy
MLP	0.511	0.512	0.504	0.748
NB	0.471	0.533	NaN	0.756
NBTree	0.536	0.522	0.527	0.748
J48	0.439	0.443	0.440	0.691
RF	0.499	0.509	0.503	0.740
Log	0.525	0.573	NaN	0.780

Tabella 1: Misure di valutazione relative ai diversi modelli con il metodo holdout

La tabella dei risultati contiene dei valori mancanti di F-Measure, questo è dovuto alla formula per il calcolo indicata nel paragrafo precedente. La classe $CDR = 2$ presenta una precision e una recall nulle poiché il modello non è stato capace di predire neanche un record. La divisione nella formula presenta valori nulli a numeratore e denominatore quindi il valore NaN viene restituito e il calcolo della media compromesso.

5.2 Iterated Holdout

La tecnica dell'iterated holdout consiste nell'iterare il metodo descritto nel paragrafo precedente. In tal modo si cerca di superare l'inconveniente della dipendenza dal partizionamento campionando dieci volte casualmente.

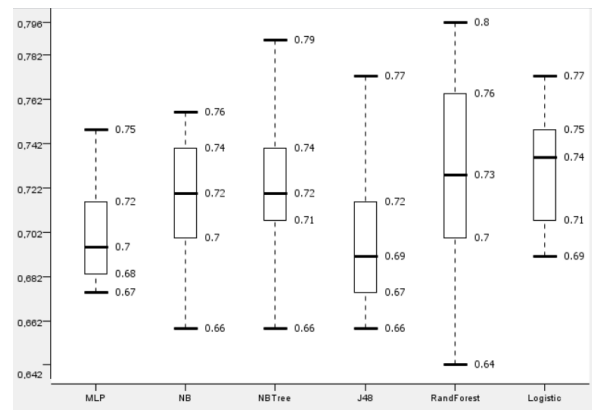


Figura 4: Il grafico mostra la variazione dell'accuracy per ogni modello

I modelli Naive Bayes tree e Random Forest si dimostrano essere quelli più sensibili al partizionamento. In generale si nota un lieve peggioramento delle performance rispetto all'Holdout, dovuto probabilmente al campionamento casuale.

Modello	Recall	Precision	F-Measure	Accuracy
MLP	0.489	0.500	0.485	0.722
NB	0.486	0.516	0.482	0.728
NBTree	0.455	0.481	0.453	0.689
J48	0.464	0.472	0.462	0.684
RF	0.492	0.506	0.498	0.733
Log	0.473	0.485	0.464	0.711

Tabella 2: Misure di valutazione relative ai diversi modelli con il metodo iterated holdout

Dalla tabella dei risultati si può notare che le accuracy sono molto prossime tra loro, senza sostanziali valori discordanti.

5.3 Cross-validation

La cross-validation utilizza un metodo leggermente diverso dall'iterated holdout. Il dataset viene suddiviso in 5 fogli di dati (o fold). L'addestramento avviene considerando un fold come *test* e

gli altri come *training* ad ogni iterazione. Le misure di performance vengono calcolate nuovamente con una media tra i vari cicli di computazione. La cross-validation permette di controllare il metodo di partizionamento, nel nostro caso abbiamo utilizzato un campionamento stratificato su CDR a causa dei valori abbastanza sbilanciati della classe target. In tal modo abbiamo potuto garantire all'interno di ogni partizione la giusta proporzione tra i vari risultati.

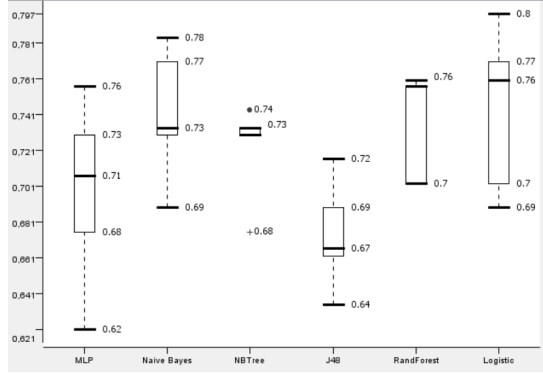


Figura 5: Il grafico mostra la misura di accuracy per ogni fold

Il grafico mostra che diversi modelli hanno ottenuto una buona accuracy, e in generale si nota una diminuzione dell'oscillazione dei risultati rispetto al modello iterated holdout. Anche in questo caso il campionamento stratificato ha avuto un ruolo importante, garantendo dei risultati mediamente simili alla tecnica dell'holdout. In particolare la tecnica di classificazione Random Forest è decisamente più stabile ed evita fenomeni di overfitting.

Modello	Recall	Precision	F-Measure	Accuracy
MLP	0.457	0.499	0.471	0.698
NB	0.498	0.541	NaN	0.741
NBTree	0.456	0.487	0.466	0.722
J48	0.433	0.455	0.441	0.674
RF	0.487	0.499	0.492	0.736
Log	0.493	0.529	Nan	0.744

Tabella 3: Misure di valutazione relative ai diversi modelli con il metodo cross validation

Dalla tabella risulta che il modello di regressione Logistica segna la accuracy migliore, sebbene poco discostante dagli altri modelli.

5.4 Feature selection

Con l'obiettivo di migliorare le performance dei modelli e l'interpretabilità degli stessi, vengono effettuate delle feature selection. Si applica un'ottimizzazione di tipo **Wrapper**: un classificatore che estrae gli attributi rilevanti. La stessa tipologia di classificatore viene utilizzata successivamente nell'addestramento; tramite questa proce-

dura vengono provati tutti i modelli definiti in precedenza. Il dataset viene suddiviso inizialmente in due sottoinsiemi *train* (67%) e *test* (33%) con campionamento stratificato su CDR. Successivamente il *train* viene suddiviso tra *training* (50%) e *validation* (50%), sempre con campionamento stratificato su CDR: queste ultime due suddivisioni vengono utilizzate dal modello wrapper per effettuare la feature selection. La scelta di dividere a metà il train set è dovuta alla scarsità di dati disponibili, un sottoinsieme troppo piccolo potrebbe portare ad errori di predizione. Operativamente si sceglie una misura di performance da ottimizzare tra: Accuracy, Precision, Recall e F-measure, si esegue la classificazione del modello tramite i nodi Feature Selection Loop Start e Feature Selection Loop End dove in quest'ultimo si dichiara la variabile da migliorare. Al termine del loop le feature vengono definitivamente scelte nel nodo Feature Selection Filter nel quale si sceglie manualmente la misura di performance migliore e con il minor numero possibile di feature, per una più facile interpretabilità. Ciascun classificatore seleziona un numero di feature diverse anche se le più presenti sono MMSE ed eTIV. Successivamente viene trainato lo stesso classificatore con il *train* set (*training* \cup *validation*) e valutato sul *test* set.

Di seguito vengono mostrate le misure di performance ottenute ottimizzando i 4 indici:

Accuracy

Modello	Recall	Precision	F-Measure	Accuracy
MLP	0.518	0.556	0.529	0.780
NB	0.471	0.520	0.479	0.740
NBTree	0.521	0.531	0.518	0.764
J48	0.521	0.531	0.518	0.764
RF	0.481	0.513	NaN	0.764
Log	0.515	0.537	0.593	0.764

Tabella 4: Operazione di feature selection per la massimizzazione della accuracy

Comparando la precedente tabella con la Tabella 1 dell'Holdout possiamo notare un generale miglioramento dell'accuracy per quasi tutti i modelli, così come le altre misure anche se non di molto.

Recall

Modello	Recall	Precision	F-Measure	Accuracy
MLP	0.545	0.558	0.549	0.764
NB	0.508	0.551	NaN	0.764
NBTree	0.505	0.530	NaN	0.764
J48	0.508	0.530	0.513	0.764
RF	0.506	0.500	0.502	0.699
Log	0.536	0.573	NaN	0.789

Tabella 5: Operazione di feature selection per la massimizzazione della recall

Nell'ottimizzazione della Recall possiamo notare come aumenti leggermente per tutti i modelli, in particolare per J48 il cui aumento è del 10%.

Precision

Modello	Recall	Precision	F-Measure	Accuracy
MLP	0.545	0.558	0.549	0.764
NB	0.478	0.536	0.490	0.748
NBTree	0.533	0.559	0.540	0.780
J48	0.495	0.510	0.495	0.724
RF	0.424	0.446	NaN	0.634
Log	0.544	0.572	0.555	0.789

Tabella 6: Operazione di feature selection per la massimizzazione della precision

Anche nell'ottimizzazione della precision possiamo osservare un leggero aumento generale. Eccezione per il Random Forest che vede una riduzione delle performance, probabilmente dovuta ad un partizionamento non favorevole, tenendo conto del fatto che il modello mostra segni di overfitting nella tecnica holdout.

F-measure

Modello	Recall	Precision	F-Measure	Accuracy
MLP	0.545	0.558	0.549	0.764
NB	0.471	0.520	0.479	0.740
NBTree	0.505	0.530	NaN	0.764
J48	0.494	0.512	0.499	0.740
RF	0.504	0.560	0.523	0.764
Log	0.544	0.572	0.555	0.789

Tabella 7: Operazione di feature selection per la massimizzazione della F-measure

Infine, l'ottimizzazione legata al F-measure risulta essere la più interessante. Oltre ad aumentare le performance dello stesso indice, migliora conseguentemente la Recall e la Precision, componenti del suo calcolo. L'accuracy, nonostante non rientri nel calcolo esplicito dell'indice F-measure, presenta anche in questo caso valori superiori al 70%, paragonabili con la Tabella 4.

6 Conclusioni

Considerando tutte le prove effettuate: Holdout, Iterated Houldout, Cross-validation e Feature Selection, il modello che si comporta meglio nella predizione della variabile di classe CDR è la **regressione logistica**. In particolare raggiunge un 79% di accuracy e 55% di f-measure medi quando si esegue la feature selection sull'ottimizzazione di quest'ultima, in particolare vengono utilizzati gli attributi di: M/F, Age, MMSE, nWBV.

Le criticità sono sorte per la distribuzione molto eterogenea del CDR nelle varie categorie e il basso numero di rilevazioni di pazienti (soltanto 373). In particolare la classe 2.0 presenta solo 3

record nel dataset in oggetto, ciò porta valori di precision e recall spesso nulli per questa classe. Sicuramente avere dati su più pazienti di questa categoria renderebbe la variabile CDR più distribuita e quindi si potrebbero addestrare modelli migliori. Per evitare di avere l'errore legato alla valutazione soggettiva di un solo medico a un solo paziente, sarebbe interessante avere a disposizione diversi CDR assegnati da medici differenti allo stesso paziente.

Infine un altro possibile miglioramento risiederebbe nel rimuovere la classe 2 del CDR, e verificare i risultati ottenuti.

Potrebbe essere interessante aggiungere attributi che identifichino la tipologia di demenza, in tal modo sarebbe possibile ricavare ulteriori informazioni predittive su determinate patologie, come l'Alzheimer. Purtroppo non è stato possibile trovare un dataset associato che fornisse questa informazione.

Riferimenti bibliografici

- [1] *Demenza* - <https://it.wikipedia.org/wiki/Demenza>
- [2] *Clinical Dementia Rating* - https://it.wikipedia.org/wiki/Clinical_Dementia_Rating
- [3] *Analysis of Human Brain MRI, RI-CHARD NORDENSKJÖLD* - <https://pdfs.semanticscholar.org/f078/58e6d1c1463367e1157f2c127b1d1a4652fd.pdf>
- [4] *MRI AND ALZHEIMERS Kaggle Dataset* - <https://www.kaggle.com/jboysen/mri-and-alzheimers>
- [5] *Socioeconomic Status (SES)* - https://www.researchgate.net/publication/11307440_Measuring_Socioeconomic_Status
- [6] *Mini-Mental State Examination* - https://it.wikipedia.org/wiki/Mini-Mental_State_Examination
- [7] *eTIV-estimated Total Intracranial Volume* - <https://surfer.nmr.mgh.harvard.edu/fswiki/eTIV>
- [8] *Atlas Scaling Factor (ASF)* - <https://www.sciencedirect.com/science/article/pii/S1053811914007769>
- [9] *CDR scala versione italiana* - https://www.demenzemedicinagenerale.net/images/test/CDR_FTL_D_Versione_Italiana_Traduzione_corretta-su-originale.pdf
- [10] *F-Measure* - <https://arxiv.org/ftp/arxiv/papers/1503/1503.06410.pdf>